Book Review

Reviewed by **Chenghao Wang**, Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University, Suzhou, China; and Department of English, School of Arts, University of Liverpool, UK,
E-mail: dancerluo@outlook.com. https://orcid.org/0009-0009-5655-3740

1 Background

Generative Artificial Intelligence (GenAI) has offered applied linguistics scholars a wealth of innovative resources, becoming widely used in language teaching, learning, testing, and research in recent years. The rise of ChatGPT has further brought GenAI into the public spotlight, with its advanced text and speech generation capabilities facilitating language skills development and gaining broad acceptance among language learners. The book *Exploring AI in Applied Linguistics* offers a comprehensive analysis of the potential and challenges posed by AI and GenAI (e.g., ChatGPT) in advancing applied linguistics. It primarily covers the evolution of AI, the capabilities of GenAI in assessment, the challenges AI presents to linguistic research and language education, and highlights directions for future GenAI research. The book comprises 15 chapters divided into four parts, and the present review evaluates each part individually and then engages critically with the findings presented in each part.

2 Introduction of the book

In Chapter 1, Chapelle et al. introduced the book by outlining its background and significance within the field of applied linguistics, particularly concerning the exploration of AI technologies. As the powerful text generation and proofreading capabilities of GenAI provide linguists and language educators with an expanded range of resources, the question arises: how can these tools be effectively and appropriately utilized in language education across different contexts and for various purposes? The following four parts offer evidence and recommendations from different perspectives.

Part I explores how various AI tools, including ChatGPT, contribute to enhancing students' language skills. In Chapter 2, Godwin-Jones et al. reviewed AI technologies

and related research, summarizing the evolution from early rule-based online machine translation to AI-integrated neural machine translation systems such as Google Translate, DeepL, and ChatGPT 3.5. The chapter also compared writing assistance tools like Grammarly and ChatGPT 3.5, highlighting GenAI's advantages in generating natural content while noting its shortcomings in corrections, individualization, and analyzing syntactic complexity and coherence. Additionally, GenAI-supported chatbots were introduced as potential partners and tutors for spoken language learning, with teachers' support highlighted as integral in AI-enhanced language education. Baumgart et al., in Chapter 3, investigated how different AI tools (Quilbot, Copy. AI and ChatGPT 3.5) influence the academic writing process. Through the analysis of students' writing texts, checklists, and interviews, it was found that the design and functionality of AI-assisted writing tools significantly influenced how students interacted with AI. Different tools were primarily used for revising, editing, developing research questions, expanding vocabulary and dialogical interaction. Although students were aware of the limitations of GenAI, many still tended to overestimate its capabilities and hold unrealistic expectations. Therefore, the authors advocate for the development of students' academic AI literacy, encouraging them to approach AI with both a critical and informed perspective. Moreover, they suggest that both the curriculum and assessment methods should be adjusted to accommodate the innovations introduced by GenAI. Kusumaningrum et al. (Chapter 4) further focused on six English learners using ChatGPT 3.5 to assist with writing emails. By analyzing both the ChatGPT's output and the learners' final email drafts, they found that learners tended to adopt AI-generated content directly without incorporating their own vocabulary or grammatical knowledge. This underscores the importance of establishing clear AI usage policies and rationales in assignment requirements.

Part II focuses on utilizing ChatGPT for assessment, including creating exam questions, grading, and providing feedback. In Chapter 5, Jia and Aryadoust examined the average scores assigned by human raters and ChatGPT 4.0 for interpreting accuracy. Their analysis demonstrates that ChatGPT effectively grasps scoring criteria, maintains consistent evaluations across diverse topics, and exhibits a moderate correlation with human ratings, thereby reliably identifying high-quality translations. However, due to the continuous updates of ChatGPT 4.0, its internal consistency is relatively poor and less stable compared to version 3.0. Additionally, they emphasized the limitation that ChatGPT 4.0 cannot analyze the delivery aspect of interpreting. Similarly, Chapter 6 examined the use of ChatGPT 4.0 for assessing content-related 74 argumentative essays, comparing its accuracy and consistency to human ratings. Kim et al. found that additional prompts and explanations do not significantly improve ChatGPT's scoring accuracy. ChatGPT 4.0 often struggles to accurately align essays with the content requirements of the writing tasks compared to manual scoring.

Subsequently, Gao et al. in Chapter 7 analyzed the effectiveness of various classifiers in detecting aberrant responses in computer-based speaking exams. The findings revealed that classifiers based on deep neural networks perform slightly better than those based on large language models (LLMs), as they incorporate audio inputs and are closely linked to speaking construct. In terms of assessment, valid and practical test questions are especially critical (Alderson et al. 1995). Chun and Barley, in Chapter 8, compared multiple-choice questions generated by GenAI and by humans, finding that the plausibility of the distractors in ChatGPT 3.5's questions still need improvement, particularly when contextual relevance is required. Focusing on GenAI in China, Xu discussed the feasibility, conceptual, and technical aspects of Chinese high school English teachers' literacy in using ErnieBot (developed by Baidu) for GenAI-based writing assessment in Chapter 9. Xu emphasized the importance of teachers' ability to analyze the quality of GenAI outputs and the necessity for teachers to independently develop their technological knowledge in a creative manner.

Building on these assessment-oriented findings, Part III shifted the focus toward prompting and integration, addressing broader research issues surrounding LLM use. Unlike the findings in Chapter 6, Xu et al. (Chapter 10) showed that by adjusting ChatGPT 4.0's temperature settings and refining the prompts, and by applying these to a corpus of 100 essays describing the qualities of a good friend, the model could more closely approximate human coding in detecting writing errors. Detailed and well-structured prompts, together with a higher temperature, markedly improved its performance. These results suggest that ChatGPT's effectiveness in evaluation may depend not only on prompt design but also on the number and type of essay samples used. Consequently, focusing on how prompting may influence GPT-4's ability to interpret pragmatic subjectivity, Su and Goslar (Chapter 11) found that GPT-4 demonstrates human-like sociopragmatic competence, namely, the ability to interpret and produce contextually appropriate language based on social norms, interpersonal relations, and levels of formality within specific sociocultural settings. They observed that, compared with general prompts, academic prompts instructing GPT-4 to "act as a linguist" significantly enhanced its sociopragmatic performance. This finding further underscores the importance of prompt framing in improving the quality and reliability of GenAI outputs. Shifting attention to AI-human interaction, Collentine (Chapter 12) incorporated LLMs into 3D games, where users engage in task-specific dialogue with virtual characters, enhancing both the learning experience and enjoyment. Collentine also highlighted the necessity of collaboration between language educators and software developers to maximize the effectiveness of AI-enabled language learning.

Part IV further explored the developments, opportunities, and challenges that AI presented to language teachers. Chapelle et al., in Chapter 13, concentrate on developing international language teachers' GenAI-based technological pedagogical

content knowledge (TPACK), a framework that emphasizes teachers' ability to integrate technology, pedagogy, and subject knowledge in coherent and effective ways. First, teachers were guided to learn how to use ChatGPT 3.5 through instructional videos, followed by engaging in creating, evaluating, and reflecting on GenAI-generated content. The questionnaires and interviews demonstrated the importance of clear instructional guidance in enhancing the user experience for teaching reading and writing. The study also highlights ChatGPT 3.5's strengths in personalizing and localizing teaching materials while acknowledging its shortcomings in terms of vocabulary level accuracy. In Chapter 14, language teachers collaborate in an immersive VR environment to complete tasks, with one participant using ChatGPT for assistance within the VR setting. Through the analysis of group discussions and post-experiment questionnaires, Compagnoni found that task-based collaboration had positive effects on both performance and the overall experience. The study also highlighted the need for prompt engineering training for teachers, the ability to design clear, purposeful, and pedagogically aligned inputs that guide GenAI systems toward producing accurate and useful outputs, as well as the importance of ensuring equitable access to these tools. Finally, Chapelle et al., in Chapter 15, summarized the book by examining two key aspects: detecting GenAI and observing its application in real-world contexts from the perspectives of researchers, learners, teachers and researcher-observers. They also advocate for further research on GenAI's language capabilities, instructional integration and anthropomorphism in specific scenarios.

3 Critical evaluation

This book makes a valuable contribution to the field of GenAI-empowered second language acquisition and language teaching. It focuses on the application of GenAI in linguistic research, providing initial explorations that offer robust methodological insights and empirical evidence for future studies and GenAI-powered language learning application development. The book serves as a resource for language learners, educators, researchers, institutions and software developers. A total of 11 chapters utilized ChatGPT (versions 3.5 and 4.0) to conduct research, revealing both its strengths and limitations. The chapters also highlight the importance of prompt engineering training for both students and teachers to enhance user experience and efficiency, thereby promoting the ecological integration of technology and pedagogy. This aligns with Rey (2025), who demonstrated that structured prompt engineering not only improved students' lexical and syntactic sophistication but also cultivated their metacognitive awareness and higher-order thinking, underscoring that prompt literacy should be systematically incorporated into teacher education programs. In

this regard, Lo's (2023) framework for prompt construction provides a valuable foundation for refining pedagogical applications of GenAI in language education.

Notably, the book mainly examines the capabilities of LLMs in text analysis and production, with particular emphasis on their relevance to the teaching of reading and writing. Only Chapters 5 and 7 address speaking-related content, and Chapter 13 briefly suggests that teachers may encourage students to use ChatGPT to generate speech scripts. In China, where there is a strong demand for English-speaking practice, many contemporary GenAI tools, integrating speech synthesis and digital human technologies, such as EAP Talk, Doubao and ErnieBot, now support both spoken and written interaction, offering considerable potential for enhancing oral English learning and instruction. Drawing on the Interaction Hypothesis (Long 1996), they can act as a native English-speaking interlocutor that engages learners in sustained conversations, thereby facilitating second language acquisition through negotiation of meaning and corrective feedback. In addition, platforms that support voice-interactive, custom-built GenAI agents (e.g., Doubao, D-ID studio) can provide personalized language practice, adaptive scaffolding, and context-sensitive feedback (Wang and Li 2025), aligning with the learner-centred pedagogical approaches increasingly adopted in language education initiatives in China. In addition, multimodal GenAI capabilities such as text-to-image generation allow teachers to visualize pedagogical content and design richer, more engaging instructional materials. Computer assisted language learning (CALL) research in China has shown growing interest in multimodal learning environments, particularly those that integrate text, audio, visuals, and conversational interaction to support learner engagement and enhance emotional experience (Wang et al. 2024). These developments resonate with the global trend toward multimodality in CALL but also demonstrate how Chinese educational contexts are advancing practical innovation at scale.

On the other hand, as most studies in this book represent preliminary explorations, there remains limited experimental evidence demonstrating whether ChatGPT can effectively enhance learners' English proficiency. Additionally, although the book acknowledges the importance of AI use guidance, for example, through instructional videos in Chapter 13, the overall discussion of how humans and AI can work together in pedagogically sound ways is still insufficient. Recent empirical work, such as Rey (2025) and Shao and Zhu (2025), consistently shows that the effectiveness of GenAI depends on well-designed human–AI collaboration rather than on AI functioning independently. These findings highlight the need for future research to move beyond descriptive accounts and adopt controlled, context-sensitive experimental designs to examine the sustained impact of GenAI on pedagogical effectiveness. Embedding ethical AI use, collaborative human–AI decision-making, and co-construction of knowledge into pedagogical frameworks will enable

educators and researchers to better realize the transformative potential of GenAI in second language education.

References

Alderson, J. Charles, Clapham Caroline & Dianne Wall. 1995. *Language test construction and evaluation*. Cambridge, UK: Cambridge University Press.

Lo, Leo S. 2023. The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship* 49(4). 102720.

Long, Michael H. 1996. The role of the linguistic environment in second language acquisition. In William Ritchie & Tej K. Bhatia (eds.), *Handbook of second language acquisition*, 413–468. San Diego: Academic Press.

Rey, Kevin Thomas. 2025. Harnessing AI in secondary education in Chinese Hong Kong: The role of prompt engineering. *Journal of China Computer-Assisted Language Learning*. in press. https://doi.org/10.1515/jccall-2025-0016.

Shao, Xiong & Yue'e Zhu. 2025. The assistance role of LLMs and NMT in student translators' Chinese–English post-editing: Differences in workload, translation quality and user perception. *Journal of China Computer-Assisted Language Learning*. in press. https://doi.org/10.1515/jccall-2025-0014.

Wang, Chenghao, Bin Zou, Yiran Du & Zixun Wang. 2024. The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System* 127. 103533.

Wang, Chenghao & Xueyun Li. 2025. Software review: Empowering language education with D-ID Creative Reality Studio's multimodal capabilities. *International Journal of Computer-Assisted Language Learning and Teaching* 15(1). 1–11.

Bionote

Chenghao Wang
Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University, Suzhou, China
Department of English, School of Arts, University of Liverpool, UK
dancerluo@outlook.com
https://orcid.org/0009-0009-5655-3740

Chenghao Wang is a PhD student at the Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University and School of Arts, University of Liverpool. His research interests include computer-assisted language learning (CALL), AI-CALL, AIGC, and VR-enhanced language learning. His work recently appears on *System, InJAL, IJHCI and Computers & Education.*